# EVALUATING TNO HUMAN TARGET DETECTION EXPERIMENTAL RESULTS AGREEMENT WITH VARIOUS IMAGE METRICS.

G. Aviram, S. R. Rotman,
Ben-Gurion University of the Negev, Department of Electrical and Computer Engineering.
P.O. Box 653 , 84105, Beer-Sheva, Israel.Tel. (972)-7-6461518. Fax. (972)-7-6472949
E-mail: guyavi@ee.bgu.ac.il srotman@ee.bgu.ac.il

## 1.SUMMARY

An evaluation of the agreement between experimental results of human target detection performance, as obtained by TNO - Human Factors Research Institute, and various image metrics is addressed in this paper. Image metrics, such as local target from background distinctness metrics (*DOYLE* and *TARGET*), a global image complexity metric (*POE*) and a textural global / local co-occurrence matrix metric (*ICOM*), are presented and applied to the TNO image database. Good agreement, denoted by relatively high correlation levels, is found between the experimental results (search rates and probabilities of detection) and both *DOYLE* and *TARGET* local image metrics values. On the other hand, a relatively low correlation level is obtained between the experimental results and the *POE* global image metric values. Correlation values obtained using the global / local *ICOM* metric are between these extremes, as expected. These results emphasize the dominance of the target to background distinctness perceptual cue and the appropriateness of the local metrics to this kind of imagery. Furthermore, they can be used to formulate empirical classification rules that can be used to evaluate and predict human detection performance in similar cases.

**Keywords:** human target detection, clutter, contrast, texture, image metrics, psychophysical experiments.

## 2. INTRODUCTION

The Search and Target Acquisition Workshop is an excellent opportunity to deal with one of the most challenging topics in the field of human image perception. This topic is the evaluation of the relationship between image content and human target detection performance. The traditional attitude to deal with this challenge is to define image metrics to measure various kinds of perceptual cues, apply them to the natural scenes and correlate their outputs with target detection experimental results. The higher the correlation, the more appropriate is the corresponding metric. Following this attitude, much work has been done in recent years, mainly with infrared imagery, creating many image metrics of different kinds, all designed to emphasize one or more of the parameters that dominate target detectability by the human observer. Although this research greatly contributed to the understanding of human perception process, no unanimous decision upon the best metrics exists. The goal of this paper and its main contribution to the workshop is to evaluate the appropriateness of various image metrics, originally designed for infrared imagery, to imagery in the visible region of the spectrum. This task was fulfilled by applying our research methodology and analysis techniques[1,2], to the TNO image database and experimental results [3] .

The paper begins with a review section describing the four image metrics we used, and containing a brief description of their main properties as well as guide lines for computational algorithms. The next section presents comments regarding the experimental database and results. The following section deals with the relationship found between the experimental data and the metrics values, including quantitative correlation analysis. The experimental data and the image metric products for each of the tested scenes are shown in an appendix.

## 3. IMAGE METRICS

Image metrics, in general, are analytical or statistical procedures, designed to describe image properties quantitatively. An image metric can be either of global, local or combined global / local orientation. A global metric is a metric applied to the entire image, and returns a quantitative measure that represents an image property. Image properties presented by global metrics are, for example, image luminance, image intensities distribution and image clutter level. A local metric describes a property of a specific image area and is usually used to determine the target distinctness from its local surrounding. A well-known local image metric is the target to background contrast. The combined global / local metric integrates global and local image measures into one metric. The *SCR* - Signal to Clutter metric is an example of a combined global / local image metric. In this work we used four image metrics that were originally designed to evaluate human target detection performance of infrared imagery. These metrics are the global edge based clutter metric (*POE*), the local target to background distinctness metrics (*DOYLE*) and (*TARGET*) and the combined global / local texture metric (*ICOM*). As a direct consequence of the different perception processes of infrared and visible imagery by humans, we noticed degradation in the global metrics performance and improvement in the local metrics performance. Some of the reasons for that are:

1. The human visual system, usually being more experienced with visible imagery, is not easily attracted by natural changes in the image intensity even if the changes are sharp as is in the case with infrared imagery. Visible imagery, being characteristically more homogenous, we expect degradation in the global edge based clutter metric performance.
2. In most natural, military oriented, infrared images, the target to background distinctness is not a significant factor that influences the observer target detection performance. The reason for that is the fact that the target edges and inner details tend to smear out and the target looses its appearance of a man made object. Moreover, temperature

*Paper presented at the RTO SCI Workshop on "Search and Target Acquisition", held in Utrecht, The Netherlands, 21-23 June 1999, and published in RTO MP-45.*

differences are not always significant to produce high local contrast levels.

These facts are supported by psychophysical experiments[2], which show that detection ability is dominated mostly by the global clutter level. On the other hand, in the case of visible imagery, the target color and sharp well-defined edges cause the target to background contrast to dominate the human perception process, and therefore one should expect to find good agreement between the local image metrics values and target detection performance.

As for the combined global / local metric, its performance depends on the image content, which determines the predominant component (global or local).

### 3.1 Probability of Edge Metric (*POE*)

The *POE* metric is a global image metric, originally designed to quantify infrared image complexity or clutter level. The metric is designed to imitate the human visual system, based on the assumption that the human eye fixates on image edges.

The technical details of the *POE* metric are as follows:

1. A *SOBEL* or similar spatial filter is used to enhance the image edges.
2. The output image intensity values are normalized to values between 0-255.
3. The resulting image is divided into $N$ blocks of twice the apparent size of the typical target in each dimension.
4. The number of points that pass a predefined threshold $T$ in each block $i$ is counted and marked as the $POE_{i,T}$ value of the current block. The threshold $T$ was chosen empirically to be 0.7 of the average pixel value of the original image in each process block.

The total image *POE* is defined as follows

$$POE = \sqrt{\frac{1}{N}\sum_{i=1}^{N} POE_{i,T}^2} \qquad (1)$$

The *POE* metric was extensively tested with infrared imagery[2,4-6], yielding good results in quantifying image clutter levels and predicting human target detection probabilities. As mentioned above, for images taken in the visible spectral region, it is expected to be less efficient.

### 3.2 Local Distinctness Metric (*DOYLE*)

The *DOYLE* local metric is based on the target and its local background intensity values distribution. It is designed to measure the differences between means and variances of the target and the local background pixel values. The basic form of the metric is as follows

$$DOYLE = \sqrt{(\mu_t - \mu_b)^2 + K * (\sigma_t - \sigma_b)^2} \qquad (2)$$

where,

$\mu_t$ , $\mu_b$ , $\sigma_t$ , $\sigma_b$ - target and background means and variances.

$K$ - weighting coefficient.

The *DOYLE* local metric was also tested extensively with infrared imagery[2,7] and produced acceptable results.

### 3.3 Local Target Complexity Metric (*TARGET*)

The *TARGET* local metric is based on the assumption that a target is more easily detected if its contrast to the background is high, if its size is big and if it has well-defined edge points relative to the interior points. While the first two conditions are trivial, the meaning of the third is that for easily detected targets the only noticed edge points are those which define the target contour. A quantitative measure for this property is defined as the target complexity (*TC*) and is equal to the area between the edge enhanced target image cumulative distribution function and a uniform cumulative distribution function as follows

$$TC = \frac{1}{L}\sum_{i=0}^{L-1} \left| S_N(i) - P(i) \right| \qquad (3)$$

where,

TC – target complexity.

L - No. of gray levels.

$S_N$ - target cumulative distribution function.

P - uniform cumulative distribution function.

The expression above is also known as the Kolmogorov - Smirnov test.

Adding the effect of the target to background contrast and the target size (*TS*) yields the *TARGET* metric as follows

$$TARGET = TS\left(contrast + \alpha TC\right) \qquad (4)$$

where $\alpha$ is a fitting parameter ($\alpha$=0.04 in our case).

Several works[4,5] are showing detailed explanation about the metric, as well as experimental results obtained with infrared imagery.

### 3.4 Combined Global / Local Texture Metric (*ICOM*)

The *ICOM* textural metric is based on the Markov co-occurrence matrix, which contains information about both the intensity values distribution and the possible transitions among neighbor pixels in the examined image area.

The co-occurrence matrix is a square (NxN) matrix where N is the number of possible pixel intensity values that occur in the image. Each image pixel contributes to the co-occurrence matrix according to its neighbor pixels intensities distribution. For the *ICOM* metric, the co-occurrence matrix is used twice. First, the textures of image areas in the size of the target are examined, looking for those areas that contain a target-like texture. Secondly, the textures of image areas of the size of the target and its local background are examined, looking for those areas that differ from the target size area texture. An intelligent combination of the two examinations can specify a target-like area, which is distinct from its local background. Such an area is probably either the real target, or an area very similar to it, attracting the human eye, and causing the human observer to classify it as a real target. The technical details for calculating the *ICOM* metric are as follows:

1. The target co-occurrence matrix – Ct is calculated in a square window containing only target pixels.
2. A square window of a size a bit smaller than the target size is stepped over the image. For each step the co-occurrence matrix of the area captured by the window – Cw is calculated.

3. The expression $x = \sum (Ct - Cw)^2$ is calculated for each step, measuring the difference between the target texture and the examined window texture.

4. A square window of a size twice the apparent target size is stepped over the image. For each step the co-occurrence matrix of the area captured by the window – Aw is calculated.

5. The expression $y = \sum (Cw - Aw)^2$ is calculated for each step, measuring the difference between the target size area texture and the examined window texture. We denoted this measure as *LOCAL ICOM.*

6. The empirical expression $z = (1-x)^4 * y^{\frac{1}{4}}$ combining the "target-like" term (x) with the "target distinctness" term (y) for each step, defines the *ICOM* image of the original scene.

7. All pixels in the *ICOM* image exceeding a predefined threshold are counted, and the result is defined as the *ICOM* value of the original scene. The larger is this value, more target like eye attracting areas are present in the image.

Further details as well as experimental results obtained using this metric are available in previous works[8,9].

## 4. COMMENTS ON THE EXPERIMENTAL DATABASE AND RESULTS

The experimental imagery database[3] includes 44, 4096*6144 pixels color images, each containing one of possibly nine military vehicles. The 44 images denoted img0001 to img0044 were divided into 3 groups. The first group contains 18 images, where the embedded target area is larger than 5000 pixels. Images indices in this group are 5, 9-10, 12-14, 18-20, 28, 31, 34-38, 40-41. The second group contains 10 images, where the extent of the target in either direction is less than 32 pixels. Images indices in this group are 4, 6, 11, 16, 23-24, 29, 32-33, 39. The third group contains the remaining 16 images.

The *DOYLE*, *TARGET* and *POE* metrics were implemented to the images of the second and the third groups, and the *ICOM* metric was implemented only to the images of the third group. The reasons for that are:

1. We believe that the very high detection performance (Pd>=0.94 for 17 images of 18) obtained for the images of the first group was dominated solely by the target size, regardless of the target distinctness or the image texture and clutter level.
2. Large computational efforts needed for the evaluation of the *ICOM* metric, forced us to degrade the image resolution. As a consequence, the targets embedded in the images of the second group appeared very narrow causing the algorithm to be less effective.

Another issue we considered regarding the imagery database is the contribution of color to detection performance. We conducted a simple, reduced scale experiment, by presenting the same images in color and in black and white to a group of observers, and asking them whether the color information helps them to detect a target, which they could not detect in the black and white image. The results of this simple experiment were almost absolute – the color information did not play any significant role in determining probability of detection. Based on this result we eliminated color from all the images, and implemented the metrics in the same way as done for the black and white infrared imagery for which they were originally designed.

## 5. ANALYSIS OF THE EXPERIMENTAL RESULTS AND METRICS VALUES

In order to analyze the agreement between experimental results and image metrics values, we plotted the values of the target detection experimental data as a function of the image metrics products, and calculated the correlation between them.

In ideal conditions we would expect the occurrence of a linear dependence, meaning, for example, that high detection probabilities will be associated with high *DOYLE* values and low detection probabilities will be associated with low *DOYLE* values. However, in practice we expect this dependence to be more ambiguous and to follow fuzzy-type logical rules.

The experimental data contains[3], for each tested scene, the number of correct, missed and false detections. Based on these results we calculated the probability of detection *(Pd)*, defined for every image as the ratio between the number of correct detections made by 62 observers, and the maximum possible number of correct detections that can be made by these observers. Another experimental measure we calculated for this analysis is the *Search Rate* defined as 1/ (Search Time).

Examples of the relationship between the experimental measures and the image metrics values are shown in figures 1-6.

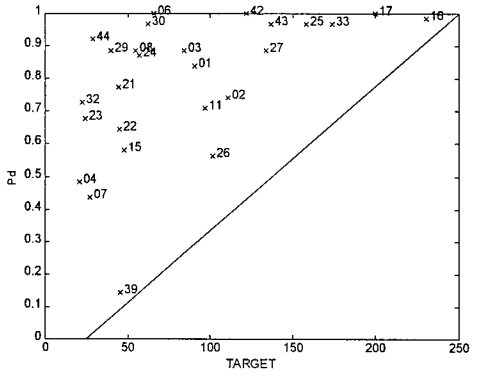


Figure 1. Pd = $f$(TARGET)

From figures 1 and 2 one can learn that high values of *DOYLE* and *TARGET* local metrics are, as expected, associated with high *Pd* values. Figures 3, 4.a and 4.b show the dependence of *Pd* upon *POE* values.

Generally speaking, it appears (fig. 3) that the global clutter level measured by *POE* is not a significant factor regarding detection performance. However, if the local target to background distinctness measured, for example, by the *DOYLE* metric is low, the image global clutter level becomes significant and dominates the detection probability (fig. 4.a). Moreover, if the *DOYLE* metric produces moderate to high values, the global clutter level has no significant role in determining the detection probability (fig. 4.b).

As for the results presented in figures 5 and 6, it appears that the scenes, which produce high *ICOM* are associated, as expected, with low value of search rate, and that the local part of the metric defined as *Local ICOM* is the dominant factor of this result.

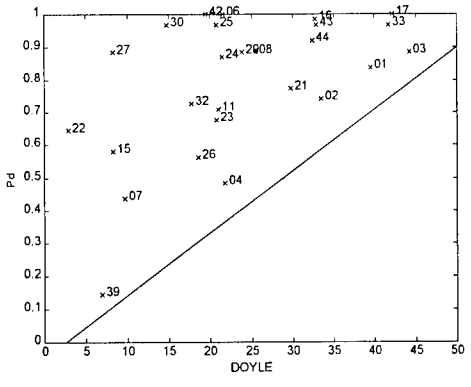


Figure 2. Pd = $f$(DOYLE)

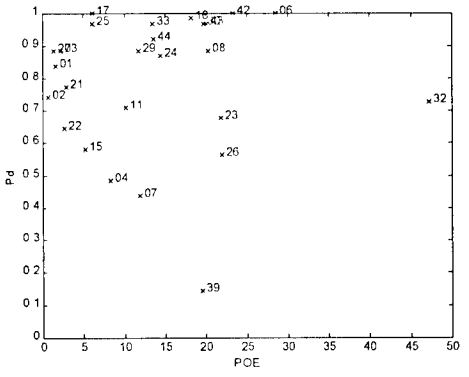


Figure 3. Pd = $f$(POE)

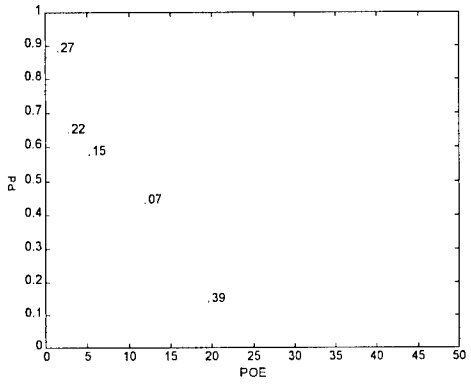


Figure 4.a Pd = $f$(POE $_{\text{LOW DOYLE}}$)

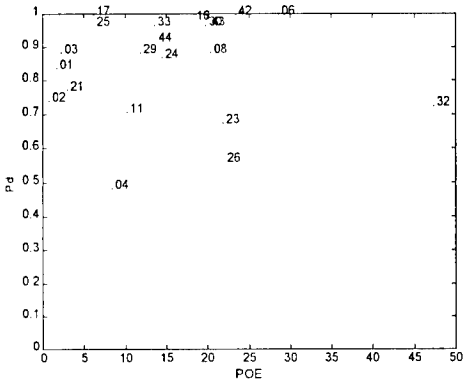


Figure 4.b Pd = $f$(POE $_{\text{MODERATE - HIGH DOYLE}}$)

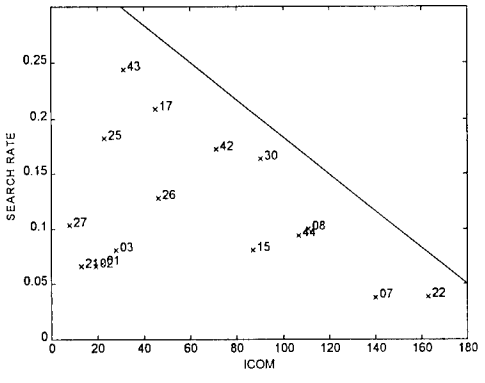


Figure 5. Search Rate = $f$(ICOM)

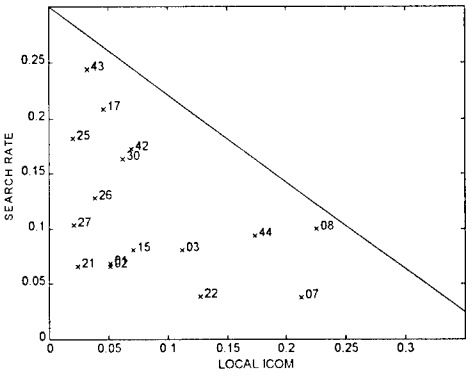


Figure 6. Search Rate = $f$(Local ICOM)

According to these figures, it also appears that the *DOYLE* and *TARGET* metrics can predict very well whether the embedded target will be easily detected, but they produce ambiguous results regarding low observable targets. The opposite occurs with *ICOM* and *Local ICOM*. This fact can be used to formulate empirical, fuzzy-type, classification rules. The fuzzy rules can lean on classification thresholds determined, for example, by the sloping line which divides the plane generated by the experimental results and the image metrics values into two regions. In order to quantify the results presented in figures 1-6, we used correlation analysis. The correlation ($\rho$) between two vectors $A$ and $B$ is defined as

$$\rho = \frac{COV(A,B)}{\sigma_A \sigma_B} = \frac{\sigma_{AB}}{\sigma_A \sigma_B} \qquad -1 \le \rho \le 1 \qquad (5)$$

where,

$\sigma_A, \sigma_B$ - standard deviation of vectors *A* and *B* respectively.

*COV(A,B)* – covariance of *A* and *B*.

In our application, the vector A contains one of the performance measures (*Search Rate* or *Pd*) as calculated for all the evaluated scenes, while vector B contains one of the image metrics values (*POE* or *DOYLE* or *TARGET or ICOM*) as calculated for the same scenes.

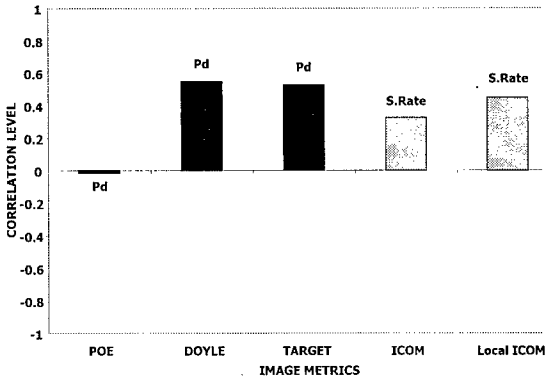


Figure 7. Correlation Values

For example, in order to calculate the correlation between the probability of detection performance factor, and the *ICOM* image metric values, we have to calculate the correlation between the two vectors *A* and *B* defined as

*A*= [Pd(Img1); Pd(Img2) ;...; Pd(Img7) ;...; Pd(Img44)]

and

*B*=[ICOM(Img1); ICOM(Img2) ;...; ICOM(Img7) ;...; ICOM(Img44)]

Representative results obtained from this analysis are shown in figure 7.

From the above figure one can learn about the relatively high correlation values obtained between both *DOYLE* and *TARGET* local metrics and *Pd*. It also appears that there is a moderate level of correlation between both *ICOM* and Local *ICOM* metrics and the *Search Rate*. Finally, there is almost no correlation between *POE* and *Pd*.

Just for the comparison, the correlation levels obtained with infrared imagery are 0.17 between the *DOYLE* values and *Pd*, 0.38 between the *POE* values and *Pd* and 0.67 between the *ICOM* values and *Search Rate* values.

## 6. CONCLUSIONS

The agreement between the TNO experimental results and various image metrics products was tested and evaluated in this paper. The analysis containing dependence and correlation measures is done after excluding images containing very large size targets (for evaluation of all the metrics), or very narrow extent targets (for evaluation of the *ICOM* metric). This is done because the experimental performance associated with these images is irrelevant in the context of the analysis methodology. The paper presents four image metrics, originally designed to evaluate detection performance of infrared imagery. These metrics have local, global and combined global / local orientation. Applying these metrics to the image database yields relatively high correlation values between the experimental results and the local metrics (*DOYLE* and *TARGET*) values. It also yields moderate correlation levels between the *ICOM* global / local texture metric values and the experimental results and very low correlation level between the *POE* global clutter metric values and the experimental results. These results emphasize that target to background distinctness is a significant perceptual cue regarding target detection in natural imagery taken in the visible region of the spectrum. Nevertheless, if the scene contains a low distinctness target, the global clutter level becomes relevant and determines the detection performance. These conclusions can be used to define fuzzy-type classification rules and to set further research directions.

## REFERENCES


1. G.Aviram, S.R.Rotman, "Human Detection Performance of Targets Embedded in Infrared Images – The Effect of Image Enhancement".

   To be published in *Optical Engineering.*

2. G.Aviram, S.R.Rotman, "Evaluation of Human Detection Performance of Targets Embedded in Natural and Enhanced Infrared Images Using Image Metrics".

   Submitted to *Optical Engineering.*

3. A.Toet, P.Bijl, F.L.Kooi, J.M.Valeton, "Image Data Set for Testing Search and Detection Models", *TNO-Human Factors Research Institute*, report TM-97-A036, (Apr. 1997).

4. S.R.Rotman, G.Tidhar, M.L.Kowalczyk, "Clutter Metrics for Target Detection Systems", *IEEE Trans. Aerospace Electron. Syst.*, Vol. 30, No. 1, pp. 81-91, (Jan. 1994).

5. G.Tidhar, G.Reiter, Z.Avital, Y.Hadar, S.R.Rotman, V.George, M.L.Kowalczyk, "Modeling Human Search and Target Acquisition Performance: IV. Detection Probability in The Cluttered Environment", *Optical Engineering*, Vol. 33, No. 3, pp. 801 - 808, (Mar. 1994).

6. S.R.Rotman, M.L.Kowalczyk, V.George, "Modeling Human Search and Target Acquisition Performance: fixation-point analysis", *Optical Engineering*, Vol. 33, No. 11, pp. 3803 - 3809, (Nov. 1994).

7. A.C.Copeland, M.M.Trivedi, J.R.McManamey, "Evaluation of Image Metrics for Target Discrimination Using Psychophysical Experiments", *Optical Engineering*, Vol. 35, No. 6, pp. 1714 - 1722, (Jun. 1996)..

8. S.R.Rotman, D.Hsu, A.Cohen, D.Shamay, M.L.Kowalczyk, "Textural Metrics for Clutter Affecting Human Target Acquisition", *Infrared Physics & Technology*, Vol. 37, pp. 667-674, (1996).

9. G.Aviram, S.R.Rotman, "Evaluating Human Detection Performance of Targets and False Alarms, in Natural and Enhanced Infrared Images, Using an Improved COM (ICOM) Textural Metric.

   To be submitted to *Optical Engineering.*

## APPENDIX

| IMAGE | Pd | Search Rate | POE | DOYLE | TARGET | ICOM |
|---|---|---|---|---|---|---|
| 1 | 0.8387 | 0.0685 | 1.6434 | 39.49 | 90.487 | 22 |
| 2 | 0.7419 | 0.0658 | 0.7298 | 33.5414 | 110.87 | 19 |
| 3 | 0.8871 | 0.0806 | 2.1878 | 44.2175 | 84.203 | 28 |
| 4 | 0.4839 | 0.0336 | 8.321 | 21.8674 | 20.979 | Narrow Extent |
| 5 | Large Size Target | | | | | |
| 6 | 1 | 0.1562 | 28.4679 | 21.5388 | 65.968 | Narrow Extent |
| 7 | 0.4355 | 0.0375 | 11.9183 | 9.7424 | 27.069 | 140 |
| 8 | 0.8871 | 0.1 | 20.3326 | 25.5205 | 54.583 | 111 |
| 9 | Large Size Target | | | | | |
| 10 | Large Size Target | | | | | |
| 11 | 0.7097 | 0.084 | 10.1294 | 20.9652 | 97.135 | Narrow Extent |
| 12 | Large Size Target | | | | | |
| 13 | Large Size Target | | | | | |
| 14 | Large Size Target | | | | | |
| 15 | 0.5806 | 0.0806 | 5.1394 | 8.313 | 47.942 | 87 |
| 16 | 0.9839 | 0.4 | 18.2353 | 32.7611 | 230.36 | Narrow Extent |
| 17 | 1 | 0.2083 | 6.117 | 42.1258 | 200 | 45 |
| 18 | Large Size Target | | | | | |
| 19 | Large Size Target | | | | | |
| 20 | Large Size Target | | | | | |
| 21 | 0.7742 | 0.0662 | 2.9254 | 29.8335 | 44.352 | 13 |
| 22 | 0.6452 | 0.0391 | 2.6516 | 2.8604 | 45.138 | 163 |
| 23 | 0.6774 | 0.0826 | 21.7671 | 20.7361 | 24.52 | Narrow Extent |
| 24 | 0.871 | 0.125 | 14.3901 | 21.4922 | 57.368 | Narrow Extent |
| 25 | 0.9677 | 0.1818 | 6.0804 | 20.8375 | 158.08 | 23 |
| 26 | 0.5645 | 0.1282 | 21.8908 | 18.6341 | 101.62 | 46 |
| 27 | 0.8871 | 0.1042 | 1.3562 | 8.281 | 133.84 | 8 |
| 28 | Large Size Target | | | | | |
| 29 | 0.8871 | 0.1333 | 11.8172 | 23.8755 | 39.795 | Narrow Extent |
| 30 | 0.9677 | 0.1639 | 19.7747 | 14.7332 | 62.26 | 90 |
| 31 | Large Size Target | | | | | |
| 32 | 0.7258 | 0.0813 | 47.2418 | 17.8229 | 22.484 | Narrow Extent |
| 33 | 0.9677 | 0.1852 | 13.5236 | 41.7339 | 173.67 | Narrow Extent |
| 34 | Large Size Target | | | | | |
| 35 | Large Size Target | | | | | |
| 36 | Large Size Target | | | | | |
| 37 | Large Size Target | | | | | |
| 38 | Large Size Target | | | | | |
| 39 | 0.1452 | 0.0287 | 19.5208 | 6.9069 | 45.353 | Narrow Extent |
| 40 | Large Size Target | | | | | |
| 41 | Large Size Target | | | | | |
| 42 | 1 | 0.1724 | 23.3694 | 19.4887 | 121.89 | 71 |
| 43 | 0.9677 | 0.2439 | 20.0798 | 32.9665 | 136.64 | 31 |
| 44 | 0.9194 | 0.0943 | 13.6477 | 32.4026 | 28.725 | 107 |

Table 1. Experimental data and image metrics products of the TNO database.